NOVEL CENTROSOME-ASSOCIATED PROTEIN AND APPLICATIONS THEREOF

The present invention relates to a novel centrosome-associated protein, to the polynucleotide encoding said protein and also to the applications of said protein and of said polynucleotide.

The cell division process consists of a nuclear division (mitosis) followed by a cytoplasmic division (cytokinesis). The mitosis is dominated by the formation of a very organized polar spindle (the mitotic spindle) consisting of two families of microtubules: polar microtubules and kinetochore microtubules. Microtubules are polymers made up of α- and β-tubulin subunits. Their growth is initiated in the peripheral region of the centrosome by a complex containing mainly a related protein, γ-tubulin. Polar microtubules are made up of rows of microtubules and of associated proteins which are put in place by the two mitotic centers, associated with centrioles, located at opposite poles of the spindle (asters). Each replicated chromosome consists of two sister chromatids connected to one another via the centromere. Kinetochore microtubules are attached to the replicated chromosomes by means of specialized structures called kinetochores, which form during prophase on each of the two faces of the centromere. The chromosomes condense during prophase and form the kinetochore microtubules, which begin to interact with the polar microtubules of the spindle after rupture of the nuclear envelope during prometaphase. Under the effect of the tension due to the opposite forces, directed toward the poles, which pull the kinetochore microtubules, the chromosomes align in the equatorial zone of the spindle during metaphase. In anaphase, under the effect of forces that are continually developed within the mitotic spindle, the sister chromatids detach and are drawn toward the

opposite poles. At the same time, the two cellular poles move apart. During telophase, the nuclear envelope re-forms at the surface of each group of chromosomes.

Cell division comes to an end when the cytoplasmic content is divided according to the process of cytokinesis. The mitotic spindle plays an important role in the process of cytokinesis, by fixing the setting up of cell segmentation. The cleavage furrow invariably appears in the plane of the equatorial plate, perpendicular to the axis of the mitotic spindle.

The processes described above are finely regulated by an equilibrium between phosphorylation reactions and dephosphorylation reactions. When the cell enters into mitosis, important changes in the phosphorylation of the proteins occur. The centrosome and the mitotic spindle are particularly enriched in phosphorylated sites. Many protein kinases, particularly serine-threonine kinases, have been described as being involved in these phosphorylation processes (in this respect, see R. Giet and C. Prigent, J. Cell Science, 112, 3591-3601, 1999). Among these, mention will be made of those located at the level of the centrosomes, among which, aurora-type kinases, that are required for centrosome separation and mitotic spindle assembly, polo-type kinases, that are involved in the maturation and formation of the bipolar spindle, and NIMA-type kinases, that regulate centrosome separation.

Mammals have at least three aurora-type protein kinases. In humans, these three protein kinases are overexpressed in cancer-related pathologies due to chromosomal anomalies. Thus, these proteins appear to play an important role in the control of ploidy. For example, inactivation or overexpression of two of these

kinases results in polyploidy. Inhibition of the activity of the aurora A kinase results in the formation of monopolar spindles. Inhibition of the activity of the aurora B kinase results in the formation of multinuclear cells through lack of cytokinesis. These chromosomal anomalies appear to be associated with disturbances in mitotic spindle formation.

The partners and the substrates of these protein kinases are still relatively unknown. For example, in xenopus, aurora A interacts with a kinesin involved in microtubule dynamics. In humans, it phosphorylates the HsTACC-3 protein, also overexpressed in many cancer cell lines. In drosophila, aurora A phosphorylates the D-TACC protein and is necessary for the localization thereof at the centrosomes in order to regulate astral microtubules. D-TACC interacts with the microtubule-associated protein (MAP) Msp, which is part of the family of XMAO215/ch-TOC/Msps proteins, which stimulate microtubule growth *in vitro* and are concentrated in the centromeres *in vivo*. D-TACC and Msp cooperate in order to stabilize centrosomes. The term "MAP" includes a collection of varied proteins defined on the basis of their ability to interact with microtubules. MAPs appear to be partners/substrates of the kinases of the centrosome, such as aurora or polo.

Correct cell division requires coordination between chromosomal segregation by the mitotic spindle and cell cleavage by the cytokinetic apparatus. The microtubules of the mitotic spindle play an essential role in both processes.

However, despite all the studies carried out on cell division, the factors that are involved in correctly setting up the mitotic spindle and/or, on the contrary, that disturb the setting up and/or the structure

thereof, thus leading to the consequences described above, are still not known.

Such knowledge would make it possible, firstly, to understand more thoroughly the mechanisms of mitosis and, secondly, to be able to develop means for combating cell division anomalies and their resulting consequences.

The present invention lies within this field.

Specifically, surprisingly and unexpectedly, the inventors have demonstrated a novel centrosome-associated human protein. By immunofluorescence, it is detected as a colocalization with the α-tubulin of the microtubules of the mitotic spindle, in particular with the aster. This protein was named ASAP, for Aster Associated Protein, by the inventors.

Overexpression of the protein according to the invention disturbs the organization of the mitotic spindle and induces aberrant and abortive mitoses (plurinuclear cells, monopolar or multipolar spindles). Its overexpression blocks cell division and, consequently, cell proliferation.

A subject of the invention is thus an isolated protein, called ASAP, characterized in that it is selected from the group consisting of:

a) a protein corresponding to the sequence represented in the attached sequence listing under the number SEQ ID NO: 1;

b) a protein exhibiting, over its entire sequence, at least 80% identity or at least 90% similarity, preferably at least 90% identity or at least 95%

similarity, with the protein of SEQ ID NO: 1.

A protein in accordance with the invention is characterized by the following properties:

- it has a molecular weight of between 60 and 100 kDa, preferably of between 65 and 80 kDa;

- it is associated with the centrosomes;

- it is colocalized, by immunofluorescence, with the α-tubulin of the microtubules of the mitotic spindle;

- it exhibits weak identity (23%) with the MAP1A protein (Microtubule Associated Protein 1A);

- it has coiled-coil domains essentially included in its C-terminal portion between, firstly, amino acids 297 and 327 and, secondly, amino acids 477 and 628, indicating either that the protein oligomerizes, or that it interacts with other proteins;

- it exhibits weak identity (20%), between amino acids 300 and 600, with a caldesmon-type domain (N.B. Gusev, Biochemistry, 10: 1112-1121, 2000), referenced pfam00769 (NCBI, domains, http://-www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid-=pfam00769), and, between amino acids 480 and 630, with a domain of ERM type (ezrin/radixin/moesin; S. Louvet-Vallet, Biol. Cell, 274: 305-316, 2000), referenced pfam02029 (NCBI, domains, http://-www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid-=pfam02029). The caldesmon and ERM proteins are also considered to be MAPs;

- it also has, between positions 65 and 303, a BRCT

domain (Breast Cancer Carboxy-Terminal domain; P. Bork et al., FASEB J., 11, 68-76 (1997)), indicating that the protein is involved in cell cycle control;

- it is very rich in α-helices in its C-terminal portion, in particular in the region between amino acids 420-620, which is almost exclusively made up of α-helices.

These elements make it possible to consider that the ASAP protein is a novel MAP.

The proteins according to the invention include any protein (natural, synthetic, semi-synthetic or recombinant) of any prokaryotic or eukaryotic organism, in particular of a mammal, comprising or consisting of an ASAP protein. Preferably, said protein is a functional ASAP protein.

The term "functional" is intended to mean a protein that has normal biological activity, i.e. that is capable of being involved in mitotic spindle organization and in cell division. This protein can comprise silent mutations that do not induce any substantial change in its activity and do not produce any phenotypic modification.

Proteins in accordance with the invention are in particular represented by the human ASAP (SEQ ID NO: 1) and murine ASAP (SEQ ID NO: 46) proteins.

Included in the proteins according to the invention defined in b) are the proteins that are variants of the sequences SEQ ID NOS: 1 and 46, in particular the proteins for which the amino acid sequence has at least one mutation corresponding in particular to a trunca-

tion, a deletion, a substitution and/or an addition of at least one amino acid residue compared with the sequences SEQ ID NOS: 1 and 46.

Preferably, the variant proteins have a mutation that results in a dysfunction (activation or inhibition) of the protein, of other genes or proteins, or else of the cell in general.

According to another advantageous embodiment of the invention, said protein is a mammalian protein, preferably a protein of human origin.

For the purpose of the present invention, the following definitions apply.

The identity of a sequence relative to the sequence of SEQ ID NO: 1 as reference sequence is assessed according to the percentage of amino acid residues that are identical, when the two sequences are aligned, so as to obtain the maximum correspondence between them.

The percentage identity can be calculated by those skilled in the art using a computer program for sequence comparison such as, for example, that of the BLAST series (Altschul et al., NAR, 1997, 25, 3389-3402).

The BLAST programs are implemented over the window of comparison consisting of the entire SEQ ID NO: 1, indicated as reference sequence.

A protein having an amino acid sequence that has at least X% identity with a reference sequence is defined, in the present invention, as a protein whose sequence can include up to 100-X alterations per 100 amino acids of the reference sequence, while at the same time

conserving the functional properties of said reference protein. For the purpose of the present invention, the term "alteration" includes consecutive or dispersed deletions, substitutions or insertions of amino acids in the reference sequence.

The similarity of a sequence relative to a reference sequence is assessed according to the percentage of amino acid residues that are identical or that differ by means of conservative substitutions, when the two sequences are aligned so as to obtain the maximum correspondence between them. For the purpose of the present invention, the term "conservative substitution" is intended to mean the substitution of an amino acid with another that has similar chemical or physical properties (size, charge or polarity), which generally does not modify the functional properties of the protein.

A protein having an amino acid sequence that has at least X% similarity with a reference sequence is defined, in the present invention, as a protein whose sequence can include up to 100-X non-conservative alterations per 100 amino acids of the reference sequence. For the purpose of the present invention, the term "non-conservative alterations" includes consecutive or dispersed non-conservative substitutions or insertions of amino acids in the reference sequence.

The expression "techniques or methods well known to those skilled in the art" is here intended to refer to the techniques or methods conventionally used by those skilled in the art and disclosed in many works, such as in particular that entitled Molecular Cloning. A Laboratory Manual (J. Sambrook, D.W. Russell (2000) Cold Spring Harbor Laboratory Press).

The protein according to the invention is obtained either from a cell, or by chemical synthesis, or by genetic recombination.

By chemical synthesis, the protein can be obtained using one of the many known peptide synthesis pathways, for example techniques using solid phases or techniques using partial solid phases, by fragment condensation or by conventional synthesis in solution. In this case, the sequence of the protein can be modified in order to improve its solubility, in particular in aqueous solvents. Such modifications are known to those skilled in the art, for instance the deletion of hydrophobic domains or the substitution of hydrophobic amino acids with hydrophilic amino acids.

The protein according to the invention consists of the series of 13 peptides corresponding to the products of translation of 13 of the 14 exons that the corresponding gene contains, the first exon not being translated (see hereinafter).

More precisely, said peptides correspond to the following sequences (positions given relative to the numbering of the sequence SEQ ID NO: 1):

- peptide 1: it comprises 25 amino acids corresponding to positions 1 to 25 (SEQ ID NO: 2);

- peptide 2: it comprises 28 amino acids corresponding to positions 26 to 53 (SEQ ID NO: 3);

- peptide 3: it comprises 107 amino acids corresponding to positions 54 to 160 (SEQ ID NO: 4);

- peptide 4: it comprises 76 amino acids corresponding to positions 161 to 236 (SEQ ID NO: 5);

- peptide 5: it comprises 31 amino acids corresponding to positions 237 to 267 (SEQ ID NO: 6);

- peptide 6: it comprises 83 amino acids corresponding to positions 268 to 350 (SEQ ID NO: 7);

- peptide 7: it comprises 24 amino acids corresponding to positions 351 to 374 (SEQ ID NO: 8);

- peptide 8: it comprises 54 amino acids corresponding to positions 375 to 428 (SEQ ID NO: 9);

- peptide 9: it comprises 32 amino acids corresponding to positions 429 to 460 (SEQ ID NO: 10);

- peptide 10: it comprises 54 amino acids corresponding to positions 461 to 514 (SEQ ID NO: 11);

- peptide 11: it comprises 49 amino acids corresponding to positions 515 to 563 (SEQ ID NO: 12);

- peptide 12: it comprises 43 amino acids corresponding to positions 564 to 606 (SEQ ID NO: 13);

- peptide 13: it comprises 41 amino acids corresponding to positions 607 to 647 (SEQ ID NO: 14).

A subject of the present invention is also a peptide consisting of a fragment of at least 10 consecutive amino acids of a protein defined above in a) or b), particularly a peptide selected from:

- the sequences corresponding to peptides 1 to 13 described above, i.e., selected from the sequences SEQ ID NO: 2 to SEQ ID NO: 14, and

- the sequences SEQ ID NOS: 47 to 53 corresponding to mutants of the hASAP protein in which there is a deletion of the N-terminal portion containing the BRCT domain (Ndel1: residues 304-647 (SEQ ID NO: 48); Ndel2: residues 411-647 (SEQ ID NO: 49); Ndel3: residues 478-647 (SEQ ID NO: 50)) or of the C-terminal portion containing the MAP domain (Cdel1: residues 1 to 477 (SEQ ID NO: 51); Cdel2: residues 1 to 418 (SEQ ID NO: 52); Cdel3: residues 1 to 303 (SEQ ID NO: 53); residues 1 to 421 (SEQ ID NO: 47)).

According to an advantageous embodiment of the invention, said peptide is useful for producing antibodies that specifically recognize a protein as defined above, preferably that recognize the ASAP protein of sequence SEQ ID NO: 1 or SEQ ID NO: 46.

The subject of the invention is thus also monoclonal or polyclonal antibodies, characterized in that they are capable of specifically recognizing a protein according to the invention.

Preferably according to the invention, the antibodies recognize, among MAPs, only and specifically the ASAP protein of sequence SEQ ID NO: 1 or SEQ ID NO: 46.

The antibodies according to the invention are, for example, chimeric antibodies, humanized antibodies, or Fab or F(ab')2 fragments. They may also be in the form of immunoconjugates or of antibodies that have been labeled in order to obtain a detectable and/or quantifiable signal.

Said antibodies can be obtained directly from human serum or from serum of animals immunized with the proteins or the peptides according to the invention.

The specific polyclonal or monoclonal antibodies can be obtained according to techniques well known to those skilled in the art.

A subject of the invention is also the use of the antibodies according to the invention, for detecting and/or purifying a protein according to the invention.

In general, the antibodies according to the invention can be advantageously used for detecting the presence of a normal or mutated protein according to the invention.

In particular, the monoclonal antibodies can be used for detecting these proteins in a biological sample. They thus constitute a means of immunocytochemical or immunohistochemical analysis of the expression of the proteins according to the invention, in particular the protein of sequence SEQ ID NO: 1, on tissue sections. In general for such analyses, the antibodies used are labeled in order to be detectable, for example by immunofluorescent compounds, by means of gold labeling, or in the form of enzymatic immunoconjugates.

They can make it possible in particular to demonstrate abnormal expression of these proteins in the biological tissues or samples, and thus allow the detection of cells exhibiting disturbances in mitotic spindle organization and/or an induction of aberrant and abortive mitoses (plurinuclear cells, monopolar or multipolar spindles) associated with overexpression of the protein according to the invention.

A subject of the invention is also a method for detecting the protein according to the invention, particularly the ASAP protein, in a biological sample, comprising a first step consisting in suitably treating

the cells by any appropriate means for making the intracellular medium accessible, a second step consisting in bringing said intracellular medium thus obtained into contact with an antibody according to the invention, and a third step consisting in demonstrating, by any appropriate means, the ASAP protein-antibody complex formed.

This method can also make it possible to measure the level of expression of the protein according to the invention in cells, particularly in cancer cells. The study of the expression of the ASAP protein (overexpression or underexpression) is an element for evaluating the proliferative capacity or the aggressiveness (ability to progress toward cancers with a poor prognosis) of cancer cells.

A subject of the invention is therefore also a method for evaluating, *in vitro*, the proliferative capacity or aggressiveness of the cancer cells contained in a biological sample, characterized in that it comprises a first step consisting in suitably treating the cells by any appropriate means for making the intracellular medium accessible, a second step consisting in bringing said intracellular medium thus obtained into contact with an antibody according to the invention, a third step consisting in demonstrating and/or measuring, by any appropriate means, the ASAP protein-antibody complex formed, and a fourth step consisting in evaluating the level of transcription of the gene, by comparing the level of ASAP protein-antibody complexes formed with that of a control biological sample selected beforehand. Said control can consist, for example, of a biological sample containing cells having a normal or altered level of proteins, to which said method is applied under the same conditions.

A subject of the invention is also a kit for carrying out any one of the methods described above, comprising:

a) at least one monoclonal or polyclonal antibody according to the invention;

b) the reagents for detecting the ASAP protein-antibody complex produced during the immunoreaction.

According to a particular embodiment of the invention, the kit can optionally comprise reagents required for making the intracellular medium accessible.

The expression "means for making the intracellular medium accessible" is intended to mean any means known to those skilled in the art, for instance cell lysis by enzymatic or chemical processes, or else sonication, membrane permeation, thermal shock.

A subject of the present invention is also an isolated polynucleotide (cDNA or genomic DNA fragment), characterized in that its sequence is selected from the group consisting of:

- the sequences encoding a protein or a peptide as defined above, and

- the sequences complementary to the preceding sequences, that may be sense or antisense.

The invention encompasses the alleles of the *asap* gene derived from any mammal, and also the polynucleotides of the natural or artificial mutants of the *asap* gene encoding an ASAP protein, particularly a functional ASAP protein as defined above.

According to an advantageous embodiment of the invention, said polynucleotide encoding an ASAP protein corresponds to a sequence selected from the group consisting of:

- the sequence SEQ ID NO: 15, corresponding to the complementary DNA of 2575 nucleotides of the mRNA encoding the human ASAP protein (hASAP);

- the sequence SEQ ID NO: 45, corresponding to the complementary DNA of 2767 nucleotides of the mRNA encoding the murine ASAP protein (mASAP);

- the genomic DNA fragment of 29750 nucleotides corresponding to the sequence represented in the attached sequence listing under the number SEQ ID NO: 16, corresponding to the human *asap* gene comprising 14 exons, only 13 of which are translated, the first exon not being translated, contained in the contig AC097467 (length 178204 base pairs) between bases 115117 and 143828 (version v.7.29a3 NCBI/Ensembl of July 12, 2002, http://www.ensembl.org), moreover located on chromosome 4q32.1 between the anonymous markers D4S1053 and D4S571 (region 161.25 megabases (Mb) to 161.28 Mb).

The sequence SEQ ID NO: 16 is contained in the BAC clone RP11-27G13 (K. Osoegawa et al., (2001) A Bacterial Artificial Chromosome Library for Sequencing the Complete Human Genome, Genome Research, Vol. 11, No. 3, 483-496, March 2001). The sequences contained in the contig AC097467 and in the BAC clone RP11-27G13 were obtained in the context of the human genome sequencing program, and have not up until now been the subject of any precise recognition or characterization making it possible to assign any function to them. Two nucleic

acids corresponding to fragments of the polynucleotide isolated by the inventors are listed in the GenBank database under the accession numbers AK024730 and AK024812, along with the ESTs listed under the accession numbers

BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952. These sequences, obtained in the context of a program of mass sequencing of human complementary DNA libraries, are incomplete and have never been either recognized or characterized. In fact, the polynucleotide isolated by the inventors exhibits long deoxyadenosine chains (poly-dA), which explains the difficulties encountered by the inventors in obtaining the complete cDNA using conventional oligo-deoxythymidine (oligo-dT) primers, said primers hybridizing randomly with the poly-dA chains. The inventors succeeded in isolating the polynucleotide corresponding to the complete mRNA by repeatedly using

the 3′ rapid amplification cDNA end (or 3′RACE) technique.

The mRNA, corresponding to the polynucleotide of sequence SEQ ID NO: 15, is specifically expressed in the testes in the form of a polynucleotide approximately 2.9 kilobases long, and in the brain in the form of a polynucleotide approximately 9 kilobases long, that may correspond either to a premessenger or to a high molecular weight isoform.

More precisely, said exons are distributed as follows on said genomic sequence (relative to the numbering of the sequence SEQ ID NO: 16):

- exon 1: it comprises 200 base pairs corresponding to positions 101 to 300 (SEQ ID NO: 17);

- exon 2: it comprises 139 base pairs corresponding to positions 1157 to 1295 (SEQ ID NO: 18);

- exon 3: it comprises 85 base pairs corresponding to positions 2050 to 2134 (SEQ ID NO: 19);

- exon 4: it comprises 321 base pairs corresponding to positions 3615 to 3935 (SEQ ID NO: 20);

- exon 5: it comprises 227 base pairs corresponding to positions 8259 to 8485 (SEQ ID NO: 21);

- exon 6: it comprises 94 base pairs corresponding to positions 14930 to 15023 (SEQ ID NO: 22);

- exon 7: it comprises 248 base pairs corresponding to positions 16715 to 16962 (SEQ ID NO: 23);

- exon 8: it comprises 71 base pairs corresponding

to positions 19552 to 19622 (SEQ ID NO: 24);

- exon 9: it comprises 169 base pairs corresponding to positions 21187 to 21355 (SEQ ID NO: 25);

- exon 10: it comprises 90 base pairs corresponding to positions 21911 to 22000 (SEQ ID NO: 26);

- exon 11: it comprises 162 base pairs corresponding to positions 23731 to 23892 (SEQ ID NO: 27);

- exon 12: it comprises 146 base pairs corresponding to positions 24014 to 24159 (SEQ ID NO: 28);

- exon 13: it comprises 133 base pairs corresponding to positions 24343 to 24475 (SEQ ID NO: 29);

- exon 14: it comprises 485 base pairs corresponding to positions 29166 to 29650 (SEQ ID NO: 30).

A subject of the invention is also:

- a fragment of any one of the polynucleotides according to the invention, of at least 15 to 1500 consecutive nucleotides, with the exclusion of the sequences listed under the accession numbers AK024730 and AK024812 and of the ESTs listed under the accession numbers BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130,

BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952 in the GenBank database, particularly a fragment selected from the sequences corresponding to the exons, i.e. selected from the sequences SEQ ID NO: 16 to SEQ ID NO: 30;

- a nucleic acid exhibiting a percentage identity of at least 80%, preferably of at least 90%, with one of the polynucleotides according to the invention.

The definition of the identity of a sequence given above for the proteins applies by analogy to the nucleic acid molecules.

Included in a polynucleotide exhibiting a percentage identity of at least 80%, preferably of at least 90%, according to the invention, are the polynucleotides that are variants of the sequences SEQ ID NOS: 15 and 45, i.e. all the polynucleotides corresponding to allelic variants, i.e. to individual variations of the sequences SEQ ID NOS: 15 and 45. These natural variant

sequences correspond to polymorphisms present in mammals, in particular in human beings, and especially to polymorphisms that may result in the occurrence of a pathology.

The term "variant polynucleotide" is also intended to denote any RNA or cDNA resulting from a mutation and/or from a variation of a splice site of the genomic sequence which has an mRNA whose complementary DNA is the polynucleotide of sequence SEQ ID NO: 15 or SEQ ID NO: 45.

Preferably, the present invention relates to the polynucleotides or the fragments that are variants of the sequences SEQ ID NOS: 15 and 45, particularly those in which the mutations result in a modification of the amino acid sequence of the proteins of sequence SEQ ID NO: 1 and SEQ ID NO: 46.

The polynucleotides according to the invention can be isolated from cells, particularly from the cells of the testes or the brain, or from cellular DNA libraries. They can also be obtained by means of a polymerase chain reaction (PCR) carried out on the total DNA of the cells or else by RT-PCR carried out on the total RNA of the cells, or by chemical synthesis.

The polynucleotides according to the invention, particularly the fragments of any one of the polynucleotides according to the invention, and the sequences listed under the accession numbers AK024730 and AK024812 and the ESTs listed under the accession numbers BU198882, BM693711,

AW372449, BM021380, BU928828, AL707573, AI885274, AI671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130, BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934,

AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952 in the GenBank database, or their fragments, can in particular be used as probes or as primers for detecting/amplifying polynucleotides (RNA or genomic DNA) corresponding to the polynucleotide according to the invention, particularly in other organisms.

The transcripts of the *asap* gene are, for example, preferably demonstrated using probes selected from the group consisting of the sequences SEQ ID NO: 15, SEQ ID NO: 45, and SEQ ID NO: 17 to SEQ ID NO: 44, or using an EST as defined above, or amplified by RT-PCR using primers selected from the group consisting of the sequences SEQ ID NOS: 31 to 43.

The polynucleotide according to the invention can make it possible to diagnose a pathological state or a genetic disease involving a dysfunction of the *asap* gene, and to screen for substances capable of modulating (activating or inhibiting) the transcription of said gene.

A subject of the invention is also the polynucleotides that can be obtained by amplification using the primers according to the invention.

The probes and primers according to the invention can be directly or indirectly labeled with a radioactive or non-radioactive compound by methods well known to those skilled in the art, in order to obtain a detectable and/or quantifiable signal.

The labeling of the probes according to the invention is carried out with radioactive elements or with non-radioactive molecules. Among the radioactive isotopes used, mention may be made of ^{32}P, ^{33}P, ^{35}S, ^{3}H or ^{125}I. The non-radioactive entities are selected from ligands such as biotin, avidin, streptavidin or digoxigenin, haptens, dyes, and luminescent agents such as radioluminescent, chemoluminescent, bioluminescent, fluorescent or phosphorescent agents.

The polynucleotides according to the invention can thus be used as a primer and/or a probe in methods using in particular the PCR (polymerase chain reaction) technique (U.S. No. 4,683,202). Other techniques for amplifying the target nucleic acid can advantageously be used as an alternative to PCR. A large number of methods currently exist that allow this amplification, for instance the SDA (Strand Displacement Amplification) technique, the TAS (Transcription-based Amplification System) technique, the 3SR (Self-sustained Sequence Replication) technique, the NASBA (Nucleic Acid Sequence Based Amplification) technique, the TMA (Transcription Mediated Amplification) technique, the LCR (Ligase Chain Reaction) technique, the RCR (Repair Chain Reaction) technique, the CPR (Cycling Probe Reaction) technique, or the Q-beta-replicase amplification technique. Mention may also be made of PCR-SSCP, which makes it possible to detect point mutations.

These techniques are of course entirely known to those

skilled in the art.

As probes or as primers, the various polynucleotides according to the invention can make it possible either to determine the transcription profile of the corresponding *asap* gene or any possible alteration of this profile in a biological sample, or to demonstrate the corresponding gene in other species, allelic variants of this gene or any possible functional alteration of this gene (substantial change in the activity of the protein encoded by said gene) resulting from a mutation (insertion, deletion or substitution) of one or more nucleotides in at least one exon of said gene. Such mutations include in particular deletions, insertions or non-conservative substitutions in codons corresponding to amino acid residues located in a domain that is essential for the biological activity of the protein.

Thus, a subject of the invention is a method for determining the transcription profile of the gene corresponding to the polynucleotide according to the invention, or an alteration in said profile, in a biological sample, comprising a first step consisting in obtaining, by any appropriate means, the total RNA from the biological sample, a second step consisting in bringing said RNA into contact with a probe according to the invention, labeled beforehand, under conventional conditions for hybridization between the RNAs and the probe, and a third step consisting in revealing, by any appropriate means, the hybrids formed.

The expression "conventional conditions for hybridization" is intended to mean those described in J. Sambrook, D.W. Russell (2000) Cold Spring Harbor Laboratory Press.

According to one embodiment of said method, the second

step can be a step consisting of reverse transcription and amplification of the transcripts, carried out using a pair of primers as described above, and the third step can be a step consisting in revealing, by any appropriate means, the amplified nucleic acids formed.

Said method for determining the transcription profile of the gene can also comprise a step consisting in evaluating the level of transcription of the gene by comparison with a control sample selected beforehand. Said control may, for example, consist of a biological sample exhibiting normal or altered transcription of the gene corresponding to the polynucleotide according to the invention, to which said method for determining the transcription profile of the gene is applied under the same conditions.

A subject of the invention is also a method for demonstrating, in other species, the gene corresponding to the polynucleotide according to the invention or the allelic variants of said gene, or a functional alteration of this gene, in a biological sample, comprising a first step consisting in obtaining, by any appropriate means, the DNA from the cells of a biological sample, a second step consisting in bringing said DNA into contact with a probe according to the invention, labeled beforehand, under conventional conditions for hybridization between the DNAs and the probe, and a third step consisting in revealing, by any appropriate means, the hybrids formed.

According to one embodiment of said method, the second step can be an amplification step carried out using a pair of primers as described above, and the third step can be a step consisting in revealing, by any appropriate means, the amplified nucleic acids formed. The method can optionally comprise a fourth step consisting

in isolating and sequencing the nucleic acids demonstrated.

A subject of the invention is also a kit of reagents for carrying out the methods described above, comprising:

a) at least one probe or one pair of primers according to the invention;

b) the reagents required for carrying out a conventional hybridization reaction between said probe or said primers and the nucleic acid of the biological sample;

c) the reagents required for carrying out an amplification reaction;

d) the reagents required for detecting and/or assaying the hybrid formed between said probe and the nucleic acid of the biological sample or the amplified nucleic acids formed.

Such a kit can also contain positive or negative controls in order to ensure the quality of the results obtained. It can also contain the reagents required for purifying the nucleic acids from the biological sample.

The polynucleotide of the invention or one of its fragments, and also the ESTs described above or their fragments, can be used to develop cell or animal models that do not express the ASAP protein, by knocking out the *ASAP* gene by means of the Si RNA method (small interfering RNA; M. McManus and P. Sharp, Nature Reviews Genetics, 3, 737-747, 2002; V. Brondani, F. Kolb, E. Billy, M/S, 6-7, 665-667, 2002) using oligonucleotides derived from their sequences.

A subject of the invention is also a cloning and/or expression vector into which the polynucleotide according to the invention is inserted.

Such a vector can contain the elements required for the expression and, optionally, the secretion of the protein in a host cell.

Said vectors preferably comprise: a promoter, translation initiation and termination signals, and also regions suitable for regulating the transcription. It should be possible for them to be maintained stably in the cell and they can optionally comprise sequences encoding specific signals specifying the secretion of the translated protein, for instance a strong ubiquitous promoter or a promoter that is selective for a particular cell and/or tissue type. These various control sequences are chosen according to the cellular host used.

The polynucleotide according to the invention can be inserted into vectors that replicate autonomously in the chosen host or vectors that are integrative with the chosen host.

Among the autonomously replicating systems, use is preferably made, according to the host cell, of systems of the plasmid or viral type. The viral vectors can in particular be adenoviruses, retroviruses, lentiviruses, poxviruses or herpesviruses. Those skilled in the art are aware of the technology that can be used for each of these systems.

When integration of the sequence into the chromosomes of the host cell is desired, use may be made, for example, of systems of the plasmid or viral type; such

viruses are, for example, retroviruses or adeno-associated viruses (AAVs).

Among the non-viral vectors, preference is given to naked polynucleotides such as naked DNA or naked RNA, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs) for expression in yeast, mouse artificial chromosomes (MACs) for expression in murine cells, and preferably human artificial chromosomes (HACs) for expression in human cells.

Such vectors are prepared according to the methods commonly used by those skilled in the art, and the recombinant vectors resulting therefrom can be introduced into the appropriate host by standard methods, for instance lipofection, electroporation, thermal shock, transformation after chemical membrane permeabilization, cell fusion.

A subject of the invention is also the transformed host cells, in particular the eukaryotic and prokaryotic cells, into which at least one polynucleotide or one fragment according to the invention or at least one vector according to the invention has been introduced.

Among the cells that can be used for the purposes of the present invention, mention may be made of bacterial cells, yeast cells, animal cells, in particular mammalian cells, or else plant cells. Mention may also be made of insect cells in which methods employing for example baculoviruses can be used.

A subject of the invention is also the nonhuman transgenic organisms, such as the transgenic animals or plants, in which all or some of the cells contain the polynucleotide according to the invention or the vector according to the invention, in free or integrated form.

Preferably according to the invention, the nonhuman transgenic organisms are those carrying cells containing a polynucleotide according to the invention, that is nonfunctional or carrying a mutation.

According to the invention, the transgenic animals are preferably mammals, except for humans, more preferably rodents, in particular mice or rats.

The transgenic animals can be obtained by any conventional method known to those skilled in the art, for instance by homologous recombination on embryonic stem cells, transfer of these stem cells to embryos, selection of the chimeras affected in the reproductive lines, and growth of such chimeras.

The transformed host cells, the transgenic animals or the transgenic plants according to the invention can thus express or overexpress the gene encoding the protein according to the invention, or their homologous gene, or express said gene into which a mutation is introduced.

The testicular or brain cells, the transformed host cells or the transgenic organisms according to the invention can be used for preparing the protein according to the invention.

The protein according to the invention, particularly the native ASAP protein, can be purified according to techniques known to those skilled in the art. Thus, the protein can be purified from cell lysates and extracts, from the culture medium supernatant, by methods used individually or in combination, such as fractionation, chromatography methods, particularly affinity chromatography methods, immunoaffinity techniques using

specific monoclonal or polyclonal antibodies, etc.

The subject of the invention is also a method for preparing the ASAP protein, characterized in that cells expressing the protein or transformed cells according to the present invention, in particular mammalian cells or the cells of transgenic organisms according to the invention, are cultured under conditions that allow the expression of said protein, and in that said protein is purified.

As a purification technique, mention may be made, for example, of affinity chromatography on glutathione-sepharose (or agarose) as described in J Sambrook & DW Russell (2000, Cold Spring Harbor Laboratory Press).

A subject of the invention is also a protein, characterized in that it can be obtained by means of any one of the methods of preparation described above.

A subject of the invention is also a method for screening for a substance capable of interacting *in vitro*, directly or indirectly, with the polynucleotide or the protein according to the invention, characterized in that:

- in a first step, the substance to be tested and the polynucleotide or the protein according to the invention are brought into contact, and
- in a second step, the complex formed between said substance and the polynucleotide or the protein according to the invention is detected by any appropriate means.

A subject of the present invention is also a method for screening for a substance capable of modulating (activating or inhibiting) the activity of the ASAP protein, characterized in that:

- in a first step, cells of a biological sample expressing the ASAP protein are brought into contact with a substance to be tested,

- in a second step, the effect of said substance on the activity of said ASAP protein is measured by any appropriate means, and

- in a third step, substances capable of modulating said activity are selected.

For the purpose of the present invention, the expression "activity of the ASAP protein" is intended to mean both the expression of the ASAP protein or of the corresponding transcripts (mRNA), and the biological activity of said ASAP protein, for instance its effect on the organization of the mitotic spindle or the induction of aberrant or abortive mitoses.

The detection of the complex formed between said substance and the polynucleotide or the protein, or the measurement of the effect of said substance on the activity of said ASAP protein, can be carried out by conventional techniques of mRNA or protein analysis that are known in themselves; by way of nonlimiting example, mention may be made of the following techniques: RT-PCR, Northern blotting, Western blotting, RIA, ELISA, immunoprecipitation, immunocytochemical or immunohistochemical analysis techniques.

Advantageously, said measurement is carried out using the probes, the primers or the antibodies as defined above.

Such substances can be biological macromolecules such as, for example, a nucleic acid, a lipid, a sugar, a protein, a peptide, a protein-lipid, protein-sugar, peptide-lipid or peptide-sugar hybrid compound, a

protein or a peptide to which have been added chemical branches or else chemical molecules.

The subject of the invention is also the polynucleotide, the protein, the antibodies, the vectors or the transformed cells according to the invention, used as medicinal products.

As indicated above, the overexpression of the protein according to the invention blocks cell division and, consequently, cell proliferation. This makes it an excellent candidate for use as an anti-mitotic agent, that can be used for example in the treatment of cancer-related pathologies.

Thus, a subject of the invention is also the use of the polynucleotide, of a vector or of the protein according to the invention, in the preparation of an anti-mitotic medicinal product.

Similarly, as is also indicated above, the overexpression of the protein according to the invention disturbs the organization of the mitotic spindle and induces aberrant and abortive mitoses (plurinuclear cells, monopolar or multipolar spindles).

Thus, a subject of the invention is also the use of an antisense polynucleotide or of an antisense fragment, of an antibody, or of a vector containing an antisense oligonucleotide, according to the invention, capable of inhibiting the expression of the polynucleotide or of the protein according to the invention, in the preparation of a medicinal product intended for the treatment of pathologies associated with disturbances in mitotic spindle organization and/or induction of aberrant and abortive mitoses (plurinuclear cells, monopolar or multipolar spindles) associated with

overexpression of the protein according to the invention.

Besides the above provisions, the invention also comprises other provisions that will emerge from the following description, which refers to examples of implementation of the invention and also to the attached drawings, in which:

- Figure 1 represents the chromosomal localization and the structure of the human *asap* gene.
- Figure 2 represents the signals obtained by Northern blotting on various human tissues after hybridization with an hASAP probe.
- Figure 3 represents the results obtained:

(A) by agarose gel electrophoresis of the RT-PCR products obtained with primers corresponding to the mouse polynucleotide, which is the ortholog of the polynucleotide SEQ ID NO: 15, using various mouse tissues;

(B) after transfer of the gel, after electrophoresis, onto a membrane and hybridization with an internal mASAP probe.

- Figure 4 represents the cellular localization of the hASAP protein coupled to the green fluorescent protein (GFP) in the 3′ position or the yellow fluorescent protein (YFP) in the 5′ position, or to an MYC tag on the N-terminal side (fusion column).

The nuclei are stained with propidium iodide or with Hoechst 33286 (4A: 63× objective; 4B, 4C and 4D: 100× objective).

- Figure 5 shows the colocalization of the human ASAP protein with alpha-tubulin. Figure 5A: cellular localization of alpha-tubulin, figure 5B: localization of the ASAP protein, figure 5C: superimposition of the 2 images showing the colocalization of the 2 proteins.

The following examples illustrate the invention but in

no way limit it.

EXAMPLE 1: Construction of the complete ASAP coding sequence:

The complete sequence of the cDNA of the ASAP protein is amplified from 2 overlapping fragments:

- a fragment A amplified by PCR from the clone AI885274 with the primers:

constFIS-1F (5'-ATGTCTGATGAAGTTTTTAGCACC-3') (SEQ ID NO: 31) and

constFIS-2R (5'-AGGCCTCAAATGATGCTAATGC-3' (SEQ ID NO: 32);

- a fragment B amplified from the clone AI671785 with the primers:

constFIS-2F (5'-ATCATTTGAGGCCTGGAAGGC-3') (SEQ ID NO: 33) and

constFIS-1R (5'-AAACACTTTTGCGAACACAGTTC-3') (SEQ ID NO: 34).

Next, in order to obtain a single PCR product corresponding to the complete sequence of the cDNA of the ASAP protein, that can be used for the function experiments, 0.5 µl of the products of each of the two PCR reactions (fragment A and B) are hybridized together at 25°C and then amplified with the primers constFIS-1F and constFIS-2F. This PCR product is subcloned into the vector PCR4 according to the producer's (Invitrogen) recommendations, and verified by sequencing.

The major difficulties encountered lay in the determination, *in silico*, of the complete ASAP coding sequence and its reconstruction *in vitro*. In particular, the choice of the primers and of the various PCRs of the 3' region were tricky due to the sequence being rich in polyA.

EXAMPLE 2: Bioinformatic analysis

Figure 1 represents the chromosomal localization and the structure of the human *asap* gene.

The complete organization of the *asap* gene and its chromosomal localization were obtained by comparing the sequence of the cDNA obtained in example 1, with the sequence of the human genome, using the Wellcome Trust Sanger Institute programs (http://www.ensembl.oriz/genome/central/ and more particularly the BLAST search program (http://genome.cse.ucsc.edu/).

The human *asap* gene consists of 29750 nucleotides comprising 14 exons, only 13 of which are translated, the first exon not being translated. The size of the exons ranges from 71 to 321 base pairs. The sequence of the gene is contained in the contig AC097467 (length 178204 base pairs) between bases 115117 and 143828 (version v.7.29a3 NCBI/Ensembl of July 12, 2002, http://www.ensembl.org) and is, moreover, located on chromosome 4q32.1 between the anonymous markers D4S1053 and D4S571 (region 161.25 megabases (Mb) to 161.28 Mb). The sequence of the gene is physically contained in the BAC clone RP11-27G13.

Two nucleic acids corresponding to fragments of the polynucleotide isolated by the inventors are listed in the GenBank database under the accession numbers AK024730 and AK024812, along with the ESTs listed under the accession numbers BU198882, BM693711, AW372449, BM021380, BU928828, AL707573, AI885274, A1671785, AA805679, BU619959, BM021126, AL598336, AW976973, BU629726, AI433877, AV751613, BQ372751, AI827535, AI866257, AA843565, R96130,

BU684090, BF958121, BQ351941, AW194906, BG203580, BF078132, AW486134, AL600279, AA025538, AL600264, BF170676, BU759494, BB025236, BF214179, AI283076, BE694273, AI266380, BM670854, AA968415, BU503982, BB700612, BE988355, BU058357, BB312934, AW061311, BM537962, BE988356, BB318982, BB311217, BB557152, BB185248, BB557128, BB698742, BB186736, AV345769, BB274293, BB632007, BB617958, AI391312, W18534, BB186581, BB311289, BB312835, AW347411, AA972439, BB263570, AU035125, BB277226, BB274224, BB268445, AW024037, AA025609, BB274174, R96089, BB272238, BB269037, BB385718, BE007324, BB325992, AJ275277, AI414381, BB125476, BB430961, BE232162, BQ121419, BQ121418, BG591509, BF457670, AL897593, AL897592, BM926692, BM538559, BI759567, AL601021, AL598780, AU222540, BG567619, AU166296, BF889835, AU164011, AV656025, BF343454, AW262441, AW237952. These sequences, obtained in the context of a program of mass sequencing of human complementary DNA libraries, are incomplete and have never been either recognized or characterized.

The protein sequence was compared to the databank sequences using the PSI-BLAST and PHI-BLAST programs of the NCBI (http://www.ncbi.nlm.nih.gov/Sitemap/). Consensus protein motifs were sought using the DART program of the NCBI and the SMART program of ExPASy-Tools (http://www.expasy.ch/tools/#similariw), the parameters of which make it possible to detect motifs with weak homology. The ASAP protein exhibits a sequence identity of 23% over the C-terminal third with a microtubule-associated protein (MAP 1A, for microtubule-associated protein 1A). Moreover, the search for conserved motifs (DART or SMART) reveals domains of caldesmon type (N.B. Gusev, Biochemistry, 10 1112-1121, 2000) and ERM type (ezrin/radixin/moesin) (Louvet-Vallet, S., Biol. Cell. 274 : 305-316, 2000), which are proteins that are also considered to be MAPs, with identities of approximately 20%. It also has a

BRCT domain (breast cancer carboxy-terminal domain; P. Bork et al., J. FASEB, 11, 68-76 (1997)) between positions 65 and 303.

The ASAP protein has coiled-coil domains essentially included in its C-terminal portion between, firstly, amino acids 297 and 327 and, secondly, amino acids 477 and 628, indicating either that the protein oligomerizes, or that it interacts with other proteins.

Computer analysis of the protein using the programs accessible on the Internet site (http://npsa-bil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_secons.html) reveals that it lacks β-sheets and is very rich in α-helices, in particular in the region between amino acids 420-620, which is almost exclusively made up of α-helices.

These elements make it possible to consider that the ASAP protein is a novel MAP.

EXAMPLE 3: Tissue expression

a) Analysis by Northern blotting

Preparation of radioactive probes:

The DNAs to be radiolabeled are isolated on a low melting point (LMP) gel according to the technique described by S. Rouquier et al. (Genomics, 17, 330-340, (1993)). Approximately 100 ng of DNA thus isolated are labeled by random priming (Klenow fragment, Proméga) in the presence of [α^{32}P dCTP] (Amersham) according to the technique described in A.P. Feinberg & B. Vogelstein (Anal. Biochem., 132, 6-13, (1983)). These probes are purified on Sephadex G-50 columns according to the technique described in J Sambrook & DW Russell (2000, Cold Spring Harbor Laboratory Press). The hybridizations are carried out overnight in the

presence of $2 \cdot 10^6$ Cpm/ml of denatured radioactive probe.

a.1) Hybridization

Two Northern blotting membranes from the company Clontech (Human MTN Blot and Human MTN Blot II, Ref. 7760-1 and 7759-1) containing human mRNAs from various tissues were hybridized with the complete hASAP cDNA labeled as described above. The membrane was hybridized in the presence of formamide at 42°C, according to the Clontech protocol. A membrane hybridization control was carried out with an actin probe. The membrane was rinsed twice at high stringency in 0.1× SSC/0.1% SDS at a temperature of 42°C, for 15 minutes. The membranes were then analyzed by autoradiography or on a PhosphorImager.

The tissues tested were: spleen, thymus, prostate, testes, ovary, small intestine, colon, blood leukocytes, heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas.

a.2) Results

Figure 2 illustrates these results.

Two signals were detected:

- a signal in the testes at approximately 2.6 kb, which corresponds to the size of the mRNA;
- a signal in the brain, but at a high molecular weight (9 kb), which corresponds either to a premessenger, or to a high molecular weight isoform.

b) Analysis by RT-PCR

This analysis was carried out on total RNA from various mouse tissues, namely brain, heart, colon, liver, small intestine, skeletal muscle, pancreas, lung, kidney, spleen and testes.

b.1) Obtaining the mouse orthologous cDNA

The total RNA from cells of various mouse tissues was extracted with the "mammalian total RNA kit" from the company Sigma. The RNAs were reverse-transcribed with the Superscript II kit from the company Invitrogen according to the conditions recommended by the supplier, and using oligodT primers. The products obtained were verified by 1% agarose gel electrophoresis. 1 µl of each sample thus obtained was, in turn, amplified by PCR (25 µl of reaction medium, 30 cycles (94°C for 15 seconds, 55°C for 30 seconds, 72°C for 30 seconds)) with primers specific for the mouse *asap* gene (mFIS-1F, 5'-ACA ACG AAT AAC AGA GTG TCC-3' (SEQ ID NO: 35) and mFIS-2R, 5'-ACT CCT GAT AAA CAG CTG CC-3' (SEQ ID NO: 36)).

The amplified products obtained were analyzed by electrophoresis on a 1% agarose gel, stained with ethydium bromide, and their size was compared with a size marker loaded onto the gel in parallel.

After electrophoresis, the amplified products obtained were transferred by capillarity onto a charged nylon membrane, in a 1.5 M NaCl/0.5 M NaOH buffer, according to the Southern technique (alkaline transfer). The membrane was then hybridized with a radiolabeled mASAP probe (SEQ ID NO: 44) generated by amplification of the sequence contained in the mouse clone AW06131 selected after comparison of the human ASAP sequence in the databanks (GenBank) (http://expression.gnf.org/-promoter/tissue/images/41739_s_at.png).

The amplification was carried out by PCR (conditions as described above, in which the reaction volume was 50 µl and the cold dCTP was at a concentration of 10 µM supplemented with 50 µCi of α-P^{32}-dCTP at 3000 Ci/mmole), using the primers mFIS-1F (SEQ ID NO: 35) and mFis-2R (SEQ ID NO: 36). The hybridizations were carried out at 65°C (in 6× SSC buffer/0.5% SDS/5× Denhardt's solution). The membrane was rinsed at high stringency (0.1× SSC/0.1% SDS), and then analyzed by autoradiography or on a PhosphorImager.

b.2) Results

Figure 3 illustrates these results.

It is noted that a major signal is obtained in the testes and the brain, which is clearly visible on the gel (Figure 3A).

After transfer of the gel and hybridization with an internal probe, it is noted that a very weak signal is detected in the other tissues (Figure 3B).

Consequently, the mRNA encoding the mASAP protein is mainly expressed in the testes and the brain. The complete mouse cDNA, amplified by RT-PCR from the mouse testicular RNA, corresponds to the sequence SEQ ID NO: 45 and the corresponding protein (mASAP) corresponds to the sequence SEQ ID NO: 46.

EXAMPLE 4: Cellular localization

a) Subcloning of the hASAP cDNA in a eukaryotic expression vector

The hASAP cDNA obtained in example 1 was inserted into

three expression vectors:

1- into pEAK10-EGFP in phase with the green fluorescent protein (GFP) fused in the C-terminal position (vector 1) (pEAK10, vector from Edge Biosystems (distributed by Q.BIOgene, Illkirch in France) into which the EGFP protein (enhanced green fluorescent protein) has been introduced according to the reference I. Gaillard et al., Eur. J. Neurosci, 15 409-418, 2002);

2- into pEYFP-C1 in phase with the yellow fluorescent protein (YFP) fused on the N-terminal side (vector 2) (distributed by BD Biosciences Clontech);

3- into GLOMYC3-1 comprising an MYC tag on the N-terminal side (vector 3), a vector derived from the vector pcDNA3.1 (Invitrogen), into which a 5' untranslated region (5'UTR) and an MYC tag have been inserted at the *HindIII-BamHI* sites, and the 3'UTR region of globin (*SpeI-XbaI* fragment) has been inserted in the *XbaI* site.

The hASAP cDNA was amplified from its initial cloning vector (pCR4-TOPO) by PCR using the pfu Turbo high-fidelity polymerase, with primers that amplify the cDNA between the starting methionine and the last amino acid. The amplified products obtained were subcloned into the three vectors.

- Cloning in PEAK-GFP. Preparation of the DNA insert by PCR [94°C 2 min; (94°C 15 sec; 58°C 30 sec; 72°C 1 min 30 sec) 30 cycles; 72°C 3 min], using the primers
hFIS-Exp1F (5'-GCCACCATGTCTGATGAAGTTTTTAGCAC-3) (SEQ ID NO: 37) and
hFIS-Exp1R (5'-GAAACACTTTTGCGAACACAGTTC-3') (SEQ ID NO: 38).

The vector was cleaved with *EcoRV* and dephosphorylated:

10 ng of vector were used for the ligation with the DNA insert. The PCR product was phosphorylated and then purified on a high PURE PCR kit (Roche): 100 ng of insert were used for the ligation [12 h at 16°C in a final volume of 10 µl (Biolabs ligase), according to standard conditions (Sambrook and Russell)].

- Cloning in Glomyc: Preparation of the DNA insert by PCR [94°C 2 min; (94°C 15 sec; 60°C 30 sec; 72°C 1 min 30 sec) 30 cycles; 72°C 3 min], using the primers:
Glomyc-FIS1F: (5'-TAATGTCTGATGAAGTTTTTAGCACC-3') (SEQ ID NO: 39) and
Glomyc-FIS1R: (5'-TCAAAACACTTTTGCGAACACAGTTC-3') (SEQ ID NO: 40).

Cloning conditions were identical to those described for the cloning in PEAK-GFP.

- Cloning in YFP: Preparation of the DNA insert: same conditions as for Glomyc, using the primers:
YFP-FIS1F (5'-AATGTCTGATGAAGTTTTTAGCACC-3') (SEQ ID NO: 41) and
Glomyc-FIS1R (SEQ ID NO: 40) (cf. above).

Cloning conditions were identical to those described for the cloning in PEAK-GFP, the vector having been cleaved beforehand with *SmaI*.

The recombinants were analyzed by PCR using a primer for the vector and an internal primer.
PEAK-GFP: annealing at 58°C, extension 45 sec at 72°C, and standard conditions for the rest. Primers: constFIS-2F (SEQ ID NO: 33) and GFP-1R (5'-TCAGCTTGCCGTAGGTGGC-3') (SEQ ID NO: 42).
YFP: annealing 55°C for 1 min: primers: YFP-2F (5'-ATGGTCCTGCTGGAGTTCG-3') (SEQ ID NO: 43) and hFIS-Exp1R

(SEQ ID NO: 38).
Glomyc: annealing 44°C, extension 45 sec at 72°C. Primers: constFIS-2F (SEQ ID NO: 33) and SP6. The recombinants were sequenced by customer-tailored automatic sequencing using the PCR products (Genome Express, Meylan).

b) Subcloning of the hASAP cDNA in a prokaryotic expression vector

Using a strategy similar to that used in paragraph a) above, the hASAP cDNA was cloned into the vector pGEX-4T2 (AMERSHAM), so as to produce a fusion protein with GST, purifiable according to standard protocols.

c) Subcloning of the mASAP cDNA in a prokaryotic or eukaryotic expression vector

Using a strategy similar to that used in paragraph a) above the mASAP cDNA was cloned into the following vectors:

- pGEX-4T2 (AMERSHAM), so as to produce a fusion protein with GST, purifiable according to standard protocols.
- pEYFP-C1 so as to produce a fusion protein (N-terminal fusion) with the yellow fluorescent protein (YFP) detectable by direct immunofluorescence.

d) Transfection, immunofluorescence and microscopy

d.1) Materials and methods

The vectors obtained were transfected according to the calcium phosphate technique or, more routinely, using the jetPEI method (GDSP10101, Qbiogene) according to the producer's recommendations, into the following cell lines:

- PEAK (ref. 37937, Edge Biosystems (distributed

by Q.BIOgene, Illkirch in France), only for the human ASAP constructs,

- HEK-293 (ATCC (American Tissue Culture Collection) reference CRL-1573; p 53 -/- non-synchronizable), for the human and murine ASAP constructs,
- nontransformed NIH3T3 (murine ASAP constructs), and
- U-2 OS (ATCC HTB-96; p 53 +/-, synchronizable).

For vectors 1) and 2) (human and murine ASAP constructs), the localizations were determined directly by detection of the GFP or YFP fluorescence at 24h, 48h and 72h, after fixing of the cells with paraformaldehyde and staining of the nuclei either with propidium iodide or with Hoechst 33286.

For vector 3) the MYC tag was detected using an anti-MYC primary antibody distributed by TEBU (9 E10, cat.#SC-40, Santa Cruz Biotechnology, CA) and an anti-mouse IgG goat secondary antibody labeled with the fluorochrome Alexa-594 (Molecular Probes, ref. A-11032, distributed in France by Interchim, Montluçon), after fixing of the cells and permeabilization thereof with 0.1% Triton X 100. The slides were analyzed, and the images were collected on a Zeiss Axiophot microscope.

d.2) Results: Cellular localization and colocalization of the ASAP protein with alpha-tubulin

- cellular localization

Figure 4 illustrates the cellular localization of the hASAP protein overexpressed in the HEK-293 line (PI = propidium iodide).

Observation under the fluorescence microscope of the slides corresponding to the various transfections with

vectors 1), 2) and 3) shows the same types of profile: the localization of the hASAP and mASAP proteins is cytoplasmic and its fibrous profile recalls that of tubulin filaments.

Moreover, it appears that the transfected cells exhibit division deficiencies since the nuclei are always larger than in the nontransfected cells (figures 4A and 4B). In addition, some of the transfected cells appear to be plurinucleated (figure 4B). This suggests abnormal division of the transfected cells.

Finally, the mitosis of the transfected cells appears to be abnormal, in terms of both the chromosomal organization and the localization profile of the hASAP and mASAP proteins at the level of the mitotic spindle. The star-shaped localization profile of the hASAP and mASAP proteins is characteristic of the nucleation of the aster microtubules around the centrosome (figures 4C and 4D).

A similar ASAP protein localization profile is detected in the U-2 OS line (p 53 +/-) overexpressing hASAP and in the nontransformed NIH 3T3 line overexpressing mASAP; an accumulation of monopolar cells in mitosis is observed.

In addition, by synchronizing the U-2 OS cells and recovering the cell extracts at various times in the cycle, it was verified that the ASAP protein was indeed present in all the phases of the cell cycle (interphase, S, G2/M).

- <u>colocalization of the ASAP protein with alpha-tubulin</u>

Figure 5 illustrates the colocalization of the human ASAP protein with alpha-tubulin; similarly, the murine

ASAP protein colocalizes with alpha-tubulin.

Figure 5A illustrates the cellular localization of alpha-tubulin detected by immunofluorescence using an anti-tubulin antibody (Alexa-594, Molecular Probe).

Figure 5B illustrates the localization of the ASAP protein labeled with YFP (yellow fluorescent protein).

Figure 5C represents the superimposition of the 2 images, demonstrating the colocalization of the 2 proteins.

EXAMPLE 5: Production of anti-hASAP and -mASAP polyclonal antibodies

a) Antibody production

The following ASAP protein constructs were cloned into the prokaryotic expression vector pGEX 4T-2 (AMERSHAM) as described in example 4:

- whole human ASAP protein (SEQ ID NO: 1),
- human protein from which the C-terminal portion containing the potential MAP domain (residues 1 to 421, SEQ ID NO: 47) has been deleted,
- whole murine protein (SEQ ID No: 46).

The proteins were expressed in *E. coli* and purified according to standard protocols. Rabbits were then immunized with the purified ASAP proteins according to a standard protocol, and the immune sera were harvested.

b) Analysis of the reactivity of the polyclonal sera with respect to the endogenous ASAP protein

The monospecific polyclonal sera directed against the

whole hASAP protein or the hASAP protein from which the C-terminal portion containing the potential MAP domain has been deleted were tested by Western blotting and by immunofluorescence, on HEK-293 and U-2 OS cells, according to standard protocols.

By Western blotting, the monospecific polyclonal serum directed against the whole hASAP protein detected a protein having an apparent molecular weight of approximately 110 kDa corresponding to the endogenous ASAP protein, in both the HEK-293 cells and the U-2 OS cells. Under these conditions, an anti-FLAG antibody detected a protein having an equivalent molecular weight, in control HEK-293 or U-2 OS cells, transfected with a vector for expression of the hASAP protein fused with a FLAG tag.

By immunofluorescence, the monospecific polyclonal serum directed against the whole hASAP protein labeled the microtubules of the HEK-293 cells in interphase, the asters of the cells in mitosis and the microtubules of the residual body at the end of telophase.

The monospecific polyclonal serum directed against the hASAP protein from which the C-terminal portion containing the potential MAP domain had been deleted exhibited the same profile by immunofluorescence and detected a protein of approximately 110 kDa, by Western blotting.

The monospecific polyclonal serum directed against the mASAP protein was used to detect which cell types expressed ASAP and at what stage(s) of the cell cycle it was expressed, by immunofluorescence on mouse testicular sections.

EXAMPLE 6: Functional analysis of the hASAP protein

using mutants from which the N-terminal portion containing the BRCT domain or the C-terminal region containing the potential MAP domain has been deleted.

Fragments of cDNA encoding an hASAP protein from which the N-terminal portion containing the BRCT domain has been deleted (Ndel1: residues 304-647 (SEQ ID NO: 48); Ndel2: residues 411-647 (SEQ ID NO: 49); Ndel3: residues 478-647 (SEQ ID NO: 50)) or from which the C-terminal portion containing the MAP domain has been deleted (Cdel1: residues 1 to 477 (SEQ ID NO: 51); Cdel2: residues 1 to 418 (SEQ ID NO: 52); Cdel3: residues 1 to 303 (SEQ ID NO: 53)) were amplified by PCR using suitable primers, and then cloned into the expression vectors pEAK10-EGFP (C-terminal fusion with GFP) and pEYFP-C1 (N-terminal fusion with YFP) according to a protocol similar to that described in example 4.

The various constructs were transfected into the HEK-293 and U-2 OS lines, and the cellular localization of the various mutants of the hASAP protein was then analyzed as described in example 4.

It is noted that, for the same deletions, a similar profile is obtained with the construct comprising YFP in the N-terminal position or GFP in the C-terminal position.

By comparison with the whole hASAP protein, the 3 constructs from which the C-terminal portion has been deleted no longer colocalize in interphase with tubulin and no longer have a fibrous appearance; these results indicate that the deletion involves a MAP domain. In addition, no monopolar cell blocked in mitosis is observed in the cells overexpressing the mutants from which the C-terminal portion containing the MAP domain

has been deleted.

By comparison with the whole hASAP protein, the three constructs from which the N-terminal portion containing the BRCT domain has been deleted exhibit a nuclear localization in the form of loci, but some fibers colocalizing with tubulin remain in the cytoplasm.

The functional analysis of the hASAP protein is completed by experiments consisting of inactivation of the expression of the gene with interfering RNAs (iRNAs).